



**University of  
Zurich<sup>UZH</sup>**

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2014

---

## **Selective attention to phonology dynamically modulates initial encoding of auditory words within the left hemisphere**

Yoncheva, Yuliya ; Maurer, Urs ; Zevin, Jason D ; McCandliss, Bruce D

**Abstract:** Selective attention to phonology, i.e., the ability to attend to sub-syllabic units within spoken words, is a critical precursor to literacy acquisition. Recent functional magnetic resonance imaging evidence has demonstrated that a left-lateralized network of frontal, temporal, and posterior language regions, including the visual word form area, supports this skill. The current event-related potential (ERP) study investigated the temporal dynamics of selective attention to phonology during spoken word perception. We tested the hypothesis that selective attention to phonology dynamically modulates stimulus encoding by recruiting left-lateralized processes specifically while the information critical for performance is unfolding. Selective attention to phonology was captured by manipulating listening goals: skilled adult readers attended to either rhyme or melody within auditory stimulus pairs. Each pair superimposed rhyming and melodic information ensuring identical sensory stimulation. Selective attention to phonology produced distinct early and late topographic ERP effects during stimulus encoding. Data-driven source localization analyses revealed that selective attention to phonology led to significantly greater recruitment of left-lateralized posterior and extensive temporal regions, which was notably concurrent with the rhyme-relevant information within the word. Furthermore, selective attention effects were specific to auditory stimulus encoding and not observed in response to cues, arguing against the notion that they reflect sustained task setting. Collectively, these results demonstrate that selective attention to phonology dynamically engages a left-lateralized network during the critical time-period of perception for achieving phonological analysis goals. These findings suggest a key role for selective attention in on-line phonological computations. Furthermore, these findings motivate future research on the role that neural mechanisms of attention may play in phonological awareness impairments thought to underlie developmental reading disabilities.

DOI: <https://doi.org/10.1016/j.neuroimage.2014.04.006>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-97364>

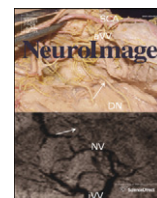
Journal Article

Published Version

Originally published at:

Yoncheva, Yuliya; Maurer, Urs; Zevin, Jason D; McCandliss, Bruce D (2014). Selective attention to phonology dynamically modulates initial encoding of auditory words within the left hemisphere. *NeuroImage*, 97:262-270.

DOI: <https://doi.org/10.1016/j.neuroimage.2014.04.006>



# Selective attention to phonology dynamically modulates initial encoding of auditory words within the left hemisphere

Yuliya Yoncheva<sup>a</sup>, Urs Maurer<sup>b</sup>, Jason D. Zevin<sup>c,d</sup>, Bruce D. McCandliss<sup>e,\*</sup>

<sup>a</sup> Department of Child and Adolescent Psychiatry, Child Study Center, New York University, USA

<sup>b</sup> Department of Psychology, University of Zurich, Switzerland

<sup>c</sup> Departments of Psychology and Linguistics, University of Southern California, USA

<sup>d</sup> Department of Linguistics, University of Southern California, USA

<sup>e</sup> Department of Psychology and Human Development, Vanderbilt University, USA

## ARTICLE INFO

### Article history:

Accepted 4 April 2014

Available online 16 April 2014

### Keywords:

Selective attention

Language

Speech

ERP

Source localization

Phonology

## ABSTRACT

Selective attention to phonology, i.e., the ability to attend to sub-syllabic units within spoken words, is a critical precursor to literacy acquisition. Recent functional magnetic resonance imaging evidence has demonstrated that a left-lateralized network of frontal, temporal, and posterior language regions, including the visual word form area, supports this skill. The current event-related potential (ERP) study investigated the temporal dynamics of selective attention to phonology during spoken word perception. We tested the hypothesis that selective attention to phonology dynamically modulates stimulus encoding by recruiting left-lateralized processes specifically while the information critical for performance is unfolding. Selective attention to phonology was captured by manipulating listening goals: skilled adult readers attended to either rhyme or melody within auditory stimulus pairs. Each pair superimposed rhyming and melodic information ensuring identical sensory stimulation. Selective attention to phonology produced distinct early and late topographic ERP effects during stimulus encoding. Data-driven source localization analyses revealed that selective attention to phonology led to significantly greater recruitment of left-lateralized posterior and extensive temporal regions, which was notably concurrent with the rhyme-relevant information within the word. Furthermore, selective attention effects were specific to auditory stimulus encoding and not observed in response to cues, arguing against the notion that they reflect sustained task setting. Collectively, these results demonstrate that selective attention to phonology dynamically engages a left-lateralized network during the critical time-period of perception for achieving phonological analysis goals. These findings suggest a key role for selective attention in on-line phonological computations. Furthermore, these findings motivate future research on the role that neural mechanisms of attention may play in phonological awareness impairments thought to underlie developmental reading disabilities.

© 2014 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-SA license (<http://creativecommons.org/licenses/by-nc-sa/3.0/>).

Phonological awareness, the ability to recognize, identify, and manipulate phonological units within a word, is central to reading acquisition (Ziegler and Goswami, 2005). Notably, tasks probing phonological awareness constitute goal-directed acoustic processing, which requires attention to certain abstract speech characteristics while ignoring other salient, yet task-irrelevant, features. Phonological skills might, therefore, rely upon selective attention mechanisms directed at abstract sub-syllabic, phonological representations (McCandliss and Yoncheva, 2011). This idea has been supported by evidence of increased blood oxygenation level-dependent (BOLD) responses within left-lateralized language-related cortical regions when literate adults selectively

attended to auditory information pertinent to rhyme judgments as opposed to competing task-irrelevant melodies (Yoncheva et al., 2010).

Recently, attentional processes have been shown to facilitate neural tracking of speech when segregating one speech stream from another speech stream (e.g., Ding and Simon, 2012; Mesgarani and Chang, 2012; Power et al., 2012). Selective attention paradigms that set up competition between channels based on low-level spatial, temporal, or spectral features have mapped out the basic auditory attentional processes (Alho et al., 2014; Fritz et al., 2007). Building on the auditory scene analysis literature (Bregman, 1990), domain-general auditory functions (e.g., event segmentation: Sridharan et al., 2007) are also being elucidated. On the other hand, how attention operates on higher-level information, such as phonological units, remains largely unknown, yet crucial to understanding the neural basis of phonological awareness.

The current study examines the temporal orchestration of the top-down mechanisms subserving selective attention to phonology. We

\* Corresponding author at: Department of Psychology and Human Development, Vanderbilt University, USA.

E-mail address: [bruce.mccandliss@vanderbilt.edu](mailto:bruce.mccandliss@vanderbilt.edu) (B.D. McCandliss).

hypothesize that left-lateralized language networks (Yoncheva et al., 2010; Joanisse et al., 2008) are recruited in a dynamic fashion as attentional processes mediating linguistic goals actively interact with the on-line encoding of spoken word stimuli. This hypothesis concerning how attention is directed to higher level linguistic representations is distinct from, yet complementary to, more basic sensory forms of selective attention, such as the notion that selective attention drives implementation of a specific task set (Dosenbach et al., 2006) engaging, for instance, a non-specific sensory amplification of any attended stimulus (Hillyard et al., 1998). In a fashion analogous to dichotic listening paradigms, which contrast event-related potential (ERP) responses to stimuli delivered to the attended versus unattended ear (Coch et al., 2005; Hillyard et al., 1973; Picton and Hillyard, 1974), we seek to delineate selective attention to phonological versus other kinds of acoustic information. Accordingly, we present auditory words simultaneously with tone melodies in a selective attention listening task. This experimental manipulation allows direct assessment on a millisecond timescale of the potential interplay between selective attention processes and spoken word encoding. We specifically expect left-lateralized networks to be dynamically recruited while the spoken word is being encoded. To characterize the putative engagement of dynamic, “bias signal” processes (Hillyard et al., 1998) as opposed to sustained ones (Dosenbach et al., 2007, 2008; Reynolds et al., 2009), a visual cue is presented prior to the auditory stimulus. Examining ERPs time-locked to the onset of this visual cue provides an opportunity to assess whether recruitment of the left-lateralized network by selective attention to phonology reflects preparatory activity for auditory stimulus perception. In contrast, examining ERPs time-locked to the onset of the combined auditory word and tone stimuli provides an opportunity to assess the impact of an adaptive, dynamic attentional mechanism operating during a time-window when linguistic information is first available for encoding.

## Materials and methods

### Participants

Sixteen right-handed monolingual native English speakers (ten female; mean age: 25.7 years, range: 20.0–39.4) took part in the study. All subjects were neurologically healthy and were screened for normal hearing, and vision. Their reading abilities ranged between the 91st and 99th percentile based on the Word Attack subtest of the Woodcock Johnson Test of Achievement (Woodcock et al., 2001). None were professional musicians. Each participant was fully briefed and provided written informed consent. Ethical approval was granted by the Institutional Review Board of the Weill Medical College of Cornell University.

### Stimuli

Stimuli were presented on a gray background (RGB = 63, 63, 63). Two visual cues were counterbalanced across attention conditions: 1) a black square enclosed in a white circle; and 2) the same square rotated by 45°. Each cue subtended 1° horizontal by 1° vertical visual angle and upon its presentation it replaced the black fixation cross (0.33° × 0.33° visual angle) that was present at the center of the screen throughout the experiment.

Auditory word/tone stimuli (Yoncheva et al., 2010, 2013) were created by simultaneous presentation of a word (mean duration = 479 ms, SD = 63 ms) and a tone triplet (total duration = 475 ms). The words were spoken by a male, native English speaker preserving their natural variability in duration and intonation. The tone triplets comprised a sequence of three unique pure tones (duration of each tone = 125 ms, silence gap between tones = 50 ms) corresponding to D, E, F#, G, A, B, or C# on the D major equal-tempered scale, and ranging in pitch from 1174.66 Hz to 2217.46 Hz. These chimeric stimulus pairs were presented over a speaker located centrally in front of the participant using E-prime 1.2 experimental control software (Psychology Software Tools,

Inc., Sharpsburg, PA). Stimulus amplitude was titrated individually for each subject to balance difficulty across the two tasks: rhyme judgment on the words and tone matching judgment on the tone triplets. Prior to the EEG session, a staircase test that progressively reduced tone amplitude, while holding word amplitude constant, was conducted to establish the stimulus amplitude level at which participants surpassed an accuracy threshold of 90% on two consecutive ten-trial sessions.

### Word selection and stimulus pairing

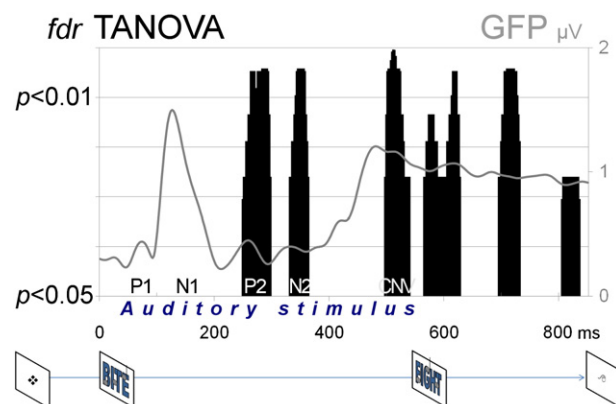
A set of 256 unique non-homophone words, each belonging to one of 32 rhyme “families” (e.g., lane, crane, stain, train) was compiled. Over the course of the experiment each word was presented twice: once as a member of a rhyming word pair and once as a member of a non-rhyming word pair. Every participant heard half of the rhyming families in the context of the rhyme task and the other half in the tone judgment task (counter-balanced across subjects).

To promote selective attention to phonology in the rhyme task, all non-rhyming trials comprised close distractors that shared either identical vowels and ended in phonologically similar consonants, (e.g., heat versus heap), or shared phonologically similar vowels and ended in identical consonants, (e.g., gum versus doom). Analogously, to promote selective attention to melody in the tone judgment task, all non-matching tone-triplets were constructed by reversing the order of the second and third tones of the triplet while maintaining the same first tone. Critically, this manipulation ensured that the disambiguating information for both the rhyming and the tone judgment task was available at approximately the same time within each trial. Finally, to ensure that rhyming decisions were based on phonological attributes rather than spelling associations, half of all rhyme targets and distractors shared spellings of rhymes and half did not.

### Procedure and task

Two tasks were performed on the pair of chimeric auditory stimuli: a rhyme judgment was performed on the words in the stimulus pair, and a tone-triplet matching judgment was performed on the tones in the stimulus pair. Contrasting the rhyme with the tone judgment task isolates the impact of selective attention to phonology.

The EEG experiment consisted of two sessions, each of which contained four blocks. A block consisted of 32 trials of the same task, with task blocks alternating within a session. Participants could take



**Fig. 1.** Top: The GFP curve (solid gray line; averaged across both tasks) illustrates the time-course of the unfolding robust auditory ERP response with the typical P1, N1, P2, and N2 components. The black vertical bars indicate time-points with whole-map topographic differences between the rhyme and tone judgment tasks (*fd*-corrected TANOVA  $p < 0.05$ ) indexing the modulation of the auditory ERP response by selective attention to phonology. Bottom: Trial sequence. Each trial began with a reminder task cue. Then a chimeric word/tone stimulus was played, followed by a second chimeric word/tone stimulus, after which participants indicated as rapidly as possible. Within the rhyme task, responses indicated whether the two auditory words rhymed. Within the tone task, responses indicated whether the two tone triplets matched.

short breaks between blocks and completed a total of 128 trials per task. The trial sequence is illustrated in Fig. 1 (bottom). Each trial within a rhyme or a tone block began with a visual task reminder cue (duration = 150 ms). After a fixed, 1500-ms interval, an auditory stimulus was presented (maximum duration of 550 ms). A second auditory stimulus was then played (SOA = 850 ms), after which participants had 1600 ms available for their two-alternative forced choice response on the relevant task. The next trial began after a normally distributed jitter of 500–1500 ms. Counter-balanced across subjects were the visual cue, prescribing the rhyme focus task (square/rotated square), and the thumb used for affirmative responses (left/right). Following this experiment the same subjects took part in a second study described elsewhere (Yoncheva et al., 2013).

#### EEG data acquisition and preprocessing

128-channel EEG was recorded using a Hydrocel Geodesic Sensor Net (Electrical Geodesics Inc., Eugene, OR) referenced to Cz (Tucker, 1993). Data were sampled at 500 Hz/channel with filters set at 0.1–200 Hz and calibrated technical zero baselines. Electrode impedances were below 50 k $\Omega$ . Data from channels with excessive artifacts were spline-interpolated, and eye blinks were corrected (multiple-source eye correction method minimizing topographic distortions (Berg and Scherg, 1994) using BESA 5.1 software). EEG data were then digitally band-pass filtered (0.1–30 Hz: 24 dB/oct, zero phase), and artifacts exceeding  $\pm 100$   $\mu$ V in any channel were automatically rejected. Correct trials were epoched from –200 to 2000 ms stimulus onset (visual reminder cue; first auditory stimulus). Single-subject potentials were averaged separately for each condition. In Brain Vision Analyzer, ERPs to the cue and the first auditory stimulus were re-referenced to average reference, then global field power (GFP; spatial root mean squared of amplitude values at all electrodes) and grand averages were computed both across tasks and separately for each task (Lehmann and Skrandies, 1980).

#### ERP analyses

To assess the effects of selective attention to phonology on the processing of spoken words, we employed an entirely data-driven strategy, carried out on two levels: scalp topographies and source localization. First, at each time point a topographic analysis of variance (TANOVA) (Strik et al., 1998) contrasting the rhyme and tone judgment tasks was performed with a two-fold purpose. Given its fine temporal precision, this approach ensured sensitive temporal localization of robust modulations by selective attention to phonology that might be invisible in *a priori* defined averaged segments (for an example of such data-driven topographic analysis of other ERP datasets see (Brem et al., 2010; Yoncheva et al., 2013)). Additionally, this temporal sensitivity for detecting stable time-windows when divergent between-task processing takes place made possible data-driven cerebral source localization and between-subject statistics. Secondly, to compare the temporal occurrence of the attentional modulations to traditional ERP components reported in the selective attention literature, segmentation was conducted based on the robust stimulus-driven ERPs, irrespective of between-task differences, and contrasts corroborated independently on a coarser temporal scale. Finally, examining ERP responses to the visual task-block reminder cue, which preceded the auditory stimulus pair, addressed the question of whether the selective attention effects were specific to the auditory word stimuli or were more consistent with a non-specific sensory amplification of any presented stimulus (Hillyard et al., 1998). Accordingly, a TANOVA at each time point was performed on ERPs to the visual reminder cue, contrasting directly the rhyme and the tone judgment tasks to capture selective attention to phonology.

Finally, to facilitate comparison with conventional ERP analyses, presented are the grand-average waveforms time-locked to the onset of

the auditory stimulus, separately for the rhyme and the tone judgment tasks. Nine non-overlapping channel clusters are created by selecting the approximate 10–10 equivalents of hallmark channels (Luu and Ferree, 2000), finding their immediate neighbors, and averaging the potentials within the cluster. The resulting clusters are: “FC5” (E28 and E20, E24, E27, E29, E34, E35); “FCz” (E6 and E5, E7, E12, E13, E106, E112); “FC6” (E117 and E110, E111, E116, E118, E123, E124); “CP5” (E47 and E41, E42, E46, E51, E52); “CPz” (E55 and Cz, E31, E54, E79, E80); “CP6” (E98 and E92, E93, E97, E102, E103); “PO7” (E65 and E58, E59, E64, E66, E69, E70); “POz” (E72 and E62, E67, E71, E76, E7); “PO8” (E90 and E83, E84, E89, E91, E95, E96).

#### Contrast of rhyme versus tone judgment tasks at each time point

TANOVA on raw (non-normalized) maps detects systematic topographic differences and overall amplitude variations between the contrasted conditions (Strik et al., 1998). TANOVA was performed on raw ERP maps (separately for the visual reminder cue and the auditory stimulus) contrasting the rhyme versus the tone judgment task for each time point in the 0–1650 ms range. To do this, global dissimilarity (the GFP of the difference map), representing a direct index of whole-map differences, was computed for each time point (Lehmann and Skrandies, 1980; Strik et al., 1998). At each time point a probability distribution was obtained via a randomization test with 5000 re-samplings. Then, a z-score of the original dissimilarity in relation to its respective distribution was computed. Multiple comparisons were accounted for by computing the local density-based false discovery rate (Strimmer, 2008a) as in our previous topographic ERP analyses (Yoncheva et al., 2013). Several statistical properties motivated its utility for our ERP data: its empirical model fitting deals with time-sample correlations inherent to the time-domain; its truncation point for model fitting minimizes false non-discovery rate (type II error) increasing leverage in interpreting both significant (auditory stimulus-locked ERP contrast) and non-significant (visual cue-locked ERP contrast) findings; the estimated local *fdr* represents the readily interpretable empirical Bayesian posterior probability of the null hypothesis (Efron, 2004, 2007). The *fdrtool* algorithm <http://strimmerlab.org/software/fdrtool/> (Strimmer, 2008b) as part of the R package archive from CRAN (R Development Core Team, 2007) was used for this analysis with input z-scores for each time-point, separately for the 825 samples of reminder cue-locked ERPs and the 825 samples of auditory stimulus-locked ERPs. The fitting parameters obtained for the reminder cue were:  $\eta_0 = 0.99$  with SD = 0.746 and for the auditory stimulus:  $\eta_0 = 0.41$  with SD = 1.659. Statistical significance was set at local *fdr*  $p < 0.05$ .

#### Source localization

We adopted a topographic mapping analysis approach, which regards multichannel EEG data as a sequence of ERP maps changing in topography and/or GFP over time (Michel, 2009; Pascual-Marqui et al., 1995). Focusing on estimating the sources underlying topographies that differ across conditions in such a data-driven manner was motivated by the axiom that different scalp topographies must have resulted from differential source contributions (Michel et al., 2004). The intracranial sources generating the topographies in these selected segments were estimated using a distributed linear inverse solution LAURA (Local A-UTO-Regressive Average (Grave de Peralta Menendez et al., 2001)) for each subject for each task. The solution space (3005 uniformly distributed points) was obtained by a SMAC procedure (Spherical Model with Anatomical Constraints (Spinelli et al., 2000)) on the Montreal Neurological Institute average 152T brain. LAURA makes no prior assumptions regarding the number of sources or their locations, can handle multiple active sources, and is thus, unlike dipole modeling, suited best for investigations of cognitive processing (Michel et al., 2004). The LAURA algorithm determines the source configuration that best simulates the biophysical behavior of the electric vector fields and provides a unique estimator of the current source density vector in the



brain; therefore estimated activity in each node depends on the activity in its neighbors in accordance with electromagnetic laws (for details and evaluation of different source estimation approaches, see (Grave de Peralta Menendez and Gonzalez Andino, 2002; Grave de Peralta Menendez et al., 2004).

TANOVA on the auditory stimulus-locked ERP contrasting rhyme with tone judgment identified six intervals with local  $fdr\ p < 0.05$ : 248–298, 330–364, 496–542, 564–628, 694–732, and 804–836 ms. For each of these intervals, potentials were averaged and sources estimated separately for the rhyme and tone judgment tasks. Paired  $t$ -tests were conducted for each node in the inverse solution space. Source estimations and statistics were implemented in Cartool 3.40 software by Denis Brunet (<https://sites.google.com/site/fbmlab/cartool>). The problem of multiple comparisons (over the 3005 source nodes) was addressed again using the unified  $fdr$  algorithm (Strimmer, 2008a), for each of the six segments separately, with input obtained from the rhyme versus tone judgment contrast  $t$ -values. The resulting fitting parameters were:  $\eta_0 = 0.50$  for the 496–542 ms segment, and  $\eta_0 = 0.43$  for the 804–836 ms segment.

#### Component segmentation

ERPs were segmented using an adaptive approach based on the minima in the GFP as markers of transitions between periods characterized by stable topographies (Lehmann and Skrandies, 1980). This data-driven identification of intervals over which to average potentials for further statistical comparisons emphasizes the robust perceptual responses that we hypothesized would be modulated by attention. Moreover, collapsing over the two tasks ensured that the selection of segments was not biased by specific between-task differences. The following segment boundaries were identified for the auditory stimulus: 44, 92, 212, 292, 360, 426, and 850 ms. For each task separately, individual's ERPs were averaged over the intervals identified above, and then contrasted directly with respect to two measures: strength of the electric field (indexed by GFP) and topographic differences across all electrodes (indexed by TANOVA on maps normalized to GFP = 1). These two complementary measures allowed complete characterization of map effects.

#### Behavioral measures

To assess processing difficulty in the rhyme and tone judgment tasks, accuracy (percent correct responses) and reaction times (RTs) for correct trials (5% trimmed means computed in SPSS 13.0 (SPSS Inc., Chicago, IL) for each condition within-subject) were analyzed.

### Results

#### Behavioral results

Both the rhyme and tone judgment tasks were performed with high accuracy. Reaction times were comparable between the two conditions (rhyme:  $M = 947.0$  ms,  $SD = 92.8$  versus tone judgment:  $M = 913.6$  ms,  $SD = 99.5$ :  $t_{15} = 1.56$ ,  $p = 0.14$ ) with a non-significant trend toward a slightly greater number of correct tone judgment trials (rhyme:  $M = 92.8\%$ ,  $SD = 4.2$  versus tone judgment:  $M = 95.9\%$ ,  $SD = 4.3$ :  $t_{15} = 1.99$ ,  $p = 0.07$ ). This general pattern suggests that the two stimulus dimensions were well equated for difficulty across the two selective attention tasks.

An additional behavioral analysis was conducted to investigate the extent to which to-be-attended versus to-be-ignored stimulus information influenced decision-making. These analyses contrasted performance on trials in which the content to be ignored led to a congruent response (i.e., rhyme and tone judgment led to the same response) with trials leading to an incongruent (opposite) response. Within each condition, this congruency analysis yielded no effects for accuracy (rhyme task:  $t_{15} = 1.59$ ,  $p = 0.13$ , tone task:  $t_{15} = 1.69$ ,  $p = 0.11$ ) or

reaction times (rhyme task:  $t_{15} = 0.18$ ,  $p = 0.86$ , tone task:  $t_{15} = 0.80$ ,  $p = 0.44$ ). Taken together these results indicate that subjects reliably attended only to the instructed stimulus dimension.

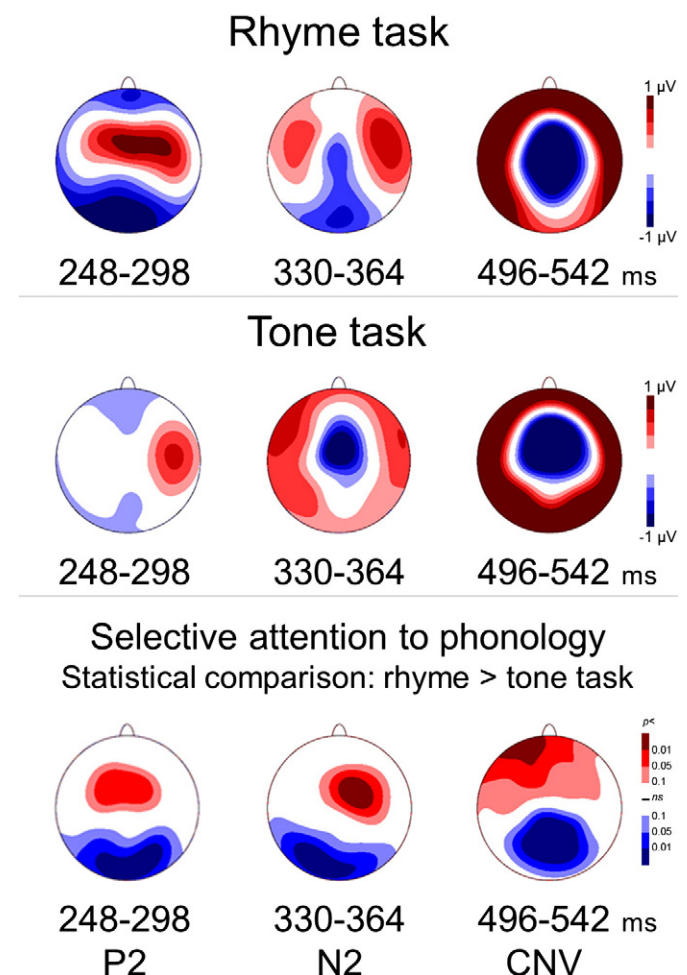
The current paradigm could be deemed relatively less challenging than our previously published paradigm (Yoncheva et al., 2013) as the former does not involve potential task switching between rhyme and tone judgment trials. Nonetheless, interference by the irrelevant stimulus dimension did not emerge suggesting that additional attentional resources were not necessarily available to support more complete processing of both the tones and rhymes within the stimulus.

#### ERP results

##### Scalp modulations by selective attention to phonology: temporal localization and characterization

TANOVAs at each time point revealed that selective attention to phonology produced robust effects in six distinct intervals: 248–298, 330–364, 496–542, 564–628, 694–732, and 804–836 ms (Fig. 1 top; Fig. 2).

Processing before and during the N1 component (92–212 ms) was unaffected by selective attention to phonology (Table 1). The first indication of divergent processing was observed in the P2 segment (212–292 ms) with both topographic and a GFP task effects reflecting a stronger and somewhat longer response during rhyming relative to



**Fig. 2.** Selective attention to phonology modulates P2, N2, and CNV components. Top two panels show voltage maps (all 129 electrodes in a planar projection) of TANOVA-defined time-windows where significant ( $fdr\ p < 0.05$ ) rhyme versus tone task differences emerged. Bottom panel shows the topography of the effect of selective attention to phonology in these averaged time-windows.

**Table 1**

Impact of selective attention to phonology on traditional auditory ERP components: TANOVA, indexing whole-map topographical differences, and Global Field Power, indexing map strength differences, between the rhyme and tone judgment tasks.

		P1 44–92 ms		N1 92–212 ms		P2 212–292 ms		N2a 292–360 ms		N2b 360–426 ms		CNV 426–850 ms	
		Mean (SEM)	Mean (SEM)	Mean (SEM)	Mean (SEM)	Mean (SEM)	Mean (SEM)	Mean (SEM)	Mean (SEM)	Mean (SEM)	Mean (SEM)	Mean (SEM)	Mean (SEM)
GFP rhyme	GFP tone	0.58 (0.04)	0.59 (0.05)	0.97 (0.16)	1.05 (0.18)	0.86 (0.09)	0.73 (0.07)	0.78 (0.07)	0.81 (0.08)	0.80 (0.07)	0.79 (0.09)	1.07 (0.09)	1.20 (0.1)
GFP p-value		ns		ns		$p < 0.05$		ns		ns		ns	
Normalized TANOVA		ns		ns		$p < 0.00001$		$p < 0.000001$		$p < 0.05$		$p < 0.000001$	

tone judgment. A robust between-condition effect during stimulus presentation emerged for the N2a segment (292–360 ms), which manifested in a topographic difference ( $p < 0.000001$ ). The final segment (426–850 ms) exhibited a CNV-like topography characterized by a central negativity/surrounding positivity, which was strongly modulated by selective attention to phonology ( $p < 0.000001$ ). The time-course of the selective attention effect is depicted also at the waveform level (Fig. 3).

Taken together, the scalp findings at the fine resolution temporal scale and the traditional component segmentation, which emphasizes the robust perceptual ERP response, converge to support the presence of significant, robust topographic modulations by selective attention to phonology during spoken word encoding. Notably, one of these modulations (496–542 ms) coincides with the final portion of the auditory word when information critical for the rhyme judgment is becoming available and is likely to be encoded and processed (Fig. 2).

#### Cerebral source contributions by selective attention to phonology during spoken word encoding

Given our *a priori* interest in isolating specifically the effect of selective attention to phonology, and in line with our previous fMRI findings (Yoncheva et al., 2010), we analyzed the sources associated with each time-window in which activity was significantly by the selective attention task manipulation. Robust task differences in the source space (local *fdr*  $p < 0.05$ ) with rhyme judgment activations greater than tone judgment activations emerged only in one segment: 496–542 ms. This contrast revealed significant source nodes, comprising a network in the left hemisphere including fusiform and lingual gyri and inferior, middle, and superior temporal regions as illustrated in Fig. 4. In contrast, the only nodes that emerged as significant from the tone > rhyme judgment contrast were in the segment that directly preceded the onset of stimulus two (804–836 ms) (Fig. 4). These consisted of bilateral frontal regions: middle frontal gyrus extending to more medial regions, as well as inferior and superior frontal gyri, pre-motor areas, and the anterior cingulate gyrus. Additionally, a right-hemispheric modulation of temporal regions (superior and mid-temporal gyrus) was observed (Fig. 4).

#### Visual task cue assay of preparatory attention

The results from the TANOVA at each time point indicated no effect of selective attention to phonology on the processing of the visual cue that served as a reminder of the current task condition (minimum  $p = 0.31$  during the entire 1650 ms duration after the reminder cue onset and before initial auditory word/tone stimulation).

## Discussion

This study investigated the impact of selective attention to phonology on ERP responses during the simultaneous presentation of two competing channels of information: one relevant to an upcoming phonological judgment, and the other, to a melodic judgment. Bottom-up stimulation was kept constant across the two tasks by using superimposed auditory stimuli. Similarly, task difficulty was equivalent between the two top-down conditions, as indicated by comparable between-task reaction time and accuracy patterns.

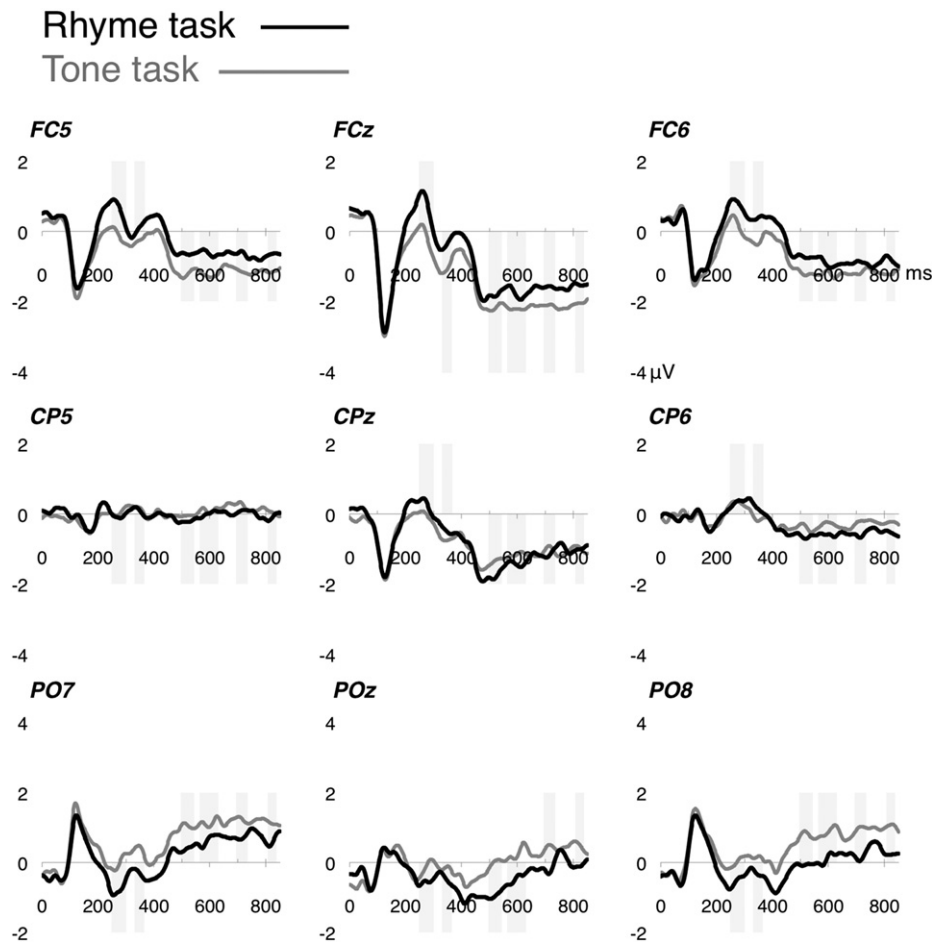
Focusing ERP analyses on the first auditory stimulus within each pair allowed examination of selective attention to phonology before decision-making and response processes could be initiated.

The current results reveal the time-course of selective attention to phonology and the engagement of the left-lateralized networks previously implicated in this process. Attentional modulations of scalp topographies emerged in distinct (early and late) time-windows during the encoding of the first word within a pair, suggesting that selective attention is recruited dynamically as linguistic information is becoming available. Given the similarities between our paradigm and classic spatial attention dichotic listening ERP paradigms that capture early modulations when selectively attending to competing spatial channels, we first discuss the early perceptual (P2/N2 component) results, which relate to early selective attention during the initial stimulus encoding. Next, we discuss the later attention effects that are relevant to understanding the mechanisms by which selective attention to phonology produce a left-lateralized modulation concurrent with the stimulus portion specifically required for phonological analysis. Finally, ERP source analysis results are examined to inform how late components may contribute to the characteristic left-lateralized cerebral signature of selectively attending to phonology, mirroring fMRI findings using the same paradigm (Yoncheva et al., 2010).

#### Early attentional modulation of encoding

Early perceptual P2/N2 responses, as sensory information was still unfolding, showed differential ERP patterns during selective attention to phonology versus melody. Capitalizing on high-density scalp coverage and data-driven time-window selection, our topographic ERP analyses allowed whole-map effect characterization beyond the *a priori* components traditionally reported in dichotic paradigms. Selective attention to phonology produced effects that were not sustained throughout the entire word but instead were present as two transient modulations of different underlying waveforms. Interestingly, both effects exhibited similar topographies: central positivity/posterior negativity based on the whole-map rhyme greater than tone judgment contrast. In light of the postulate that a stable ERP topography can reflect a particular brain state (Michel et al., 2004), the observed matching topographic patterns can be interpreted as common (linguistic) processes involved in mediating the early selective attention effects on perception.

The present modulations of the robust short-latency ERP components conform with the time-course of auditory selective attention, established by classic dichotic listening studies that contrast attention with inattention conditions on identical simple tone stimuli (Hillyard et al., 1973; Picton and Hillyard, 1974). To help constrain the interpretation of our attention to phonology effects, let us consider dichotic listening investigations of the perception of two coherent narratives while participants pursue various linguistic goals, e.g., syllable monitoring (Hink et al., 1978) or phrase-by-phrase repetition and narrative comprehension (Woods et al., 1984). ERPs to short linguistic probes superimposed on attended versus unattended prose passages have demonstrated a common profile of waveform enhancement of early perceptual processing (Hink and Hillyard, 1976; Hink et al., 1978). In



**Fig. 3.** Grand-average waveforms of the rhyme task (black) and tone task (gray) ERPs at nine channel clusters, covering approximately the 10–10 equivalents of fronto-central (FC5/FCz/FC6), centro-parietal (CP5/CPz/CP6) and occipito-parietal (PO7/POz/PO8) sites. Gray vertical rectangles illustrate significant ( $p < .05$ ) selective attention effects within the TANOVA analysis, and the valence of each corresponding ERP response within each channel.

contrast, non-linguistic (tone) probes in the same tasks have exhibited a different, or absent, modulation (Woods et al., 1984). Together these findings support the idea that the attentional selection of early components observed in the rhyme task reflect tuning to the complex features of speech, rather than a mere suppression of tone-induced activation. More recently, strides have been made in unraveling the temporal dynamics of the neural mechanisms mediating segregating one speech stream from another (e.g., Ding and Simon, 2012; Mesgarani and Chang, 2012; Power et al., 2012) typically building upon a wealth of investigations of auditory scene analysis (Bregman, 1990). Notably, the current study offers a novel perspective on the importance of selective attention to phonology in the face of competing irrelevant non-linguistic information.

#### Later modulation by left-lateralized cerebral sources

Unlike selective attention paradigms that continually direct attention to a certain channel, selection of rhyme-relevant information in this study was required within a specific, predictable time window during the first stimulus of each trial. It was indeed this time window containing the critical rhyme information that showed significantly different scalp topographies while attending to phonology versus melody, suggesting that selective attention to phonology operates in a dynamic fashion rather than merely sustaining task-set maintenance.

Auditory ERPs in linguistic tasks exhibit bottom-up sensitivity to phonological factors, even in the absence of explicit phonological analysis demands (Bonte and Blomert, 2004; Desroches et al., 2009; Praamstra and

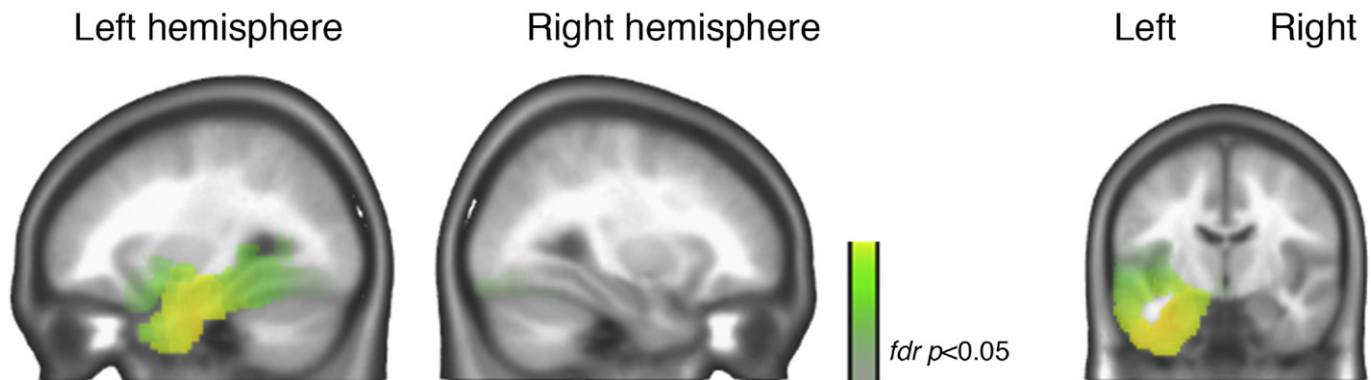
Stegeman, 1993). Crucially, such differential ERP patterns are temporally aligned with the timing of relevant stimulus properties, i.e., initial phonological overlap within a word pair elicits earlier ERP divergence than does final phonological overlap (Dumay et al., 2001; Newman and Connolly, 2009; Praamstra et al., 1994). These examples inherently fit with temporal attention's engagement in natural speech perception (Astheimer and Sanders, 2009; Sanders et al., 2002).

The ERP topography during the rhyme-relevant time window (496–542 ms) modulated by selective attention to phonology resembles that of a contingent negative variation (CNV). CNV can index domain-general activity relevant to preparing for a specific action. Neuroimaging evidence has associated CNV generation with BOLD activity in a thalamo-cortico-striatal network (Fan et al., 2007; Nagai et al., 2004). This relation dovetails with subdural recording findings that the scalp-recorded CNV represents a summation of multiple cortical potentials with different origins and functions (Hamano et al., 1997). The likely malleability of the CNV by processing goals and the coincidence of the present CNV modulation with the availability of rhyme information collectively suggest that the CNV effect is relevant to the perception for action required by the phonological analysis demands of the rhyming task. Oscillatory neural mechanisms supporting active sensing (Schroeder et al., 2010), including temporal prediction in speech perception (Schroeder et al., 2008), are candidate mediators of our attentional selection effect.

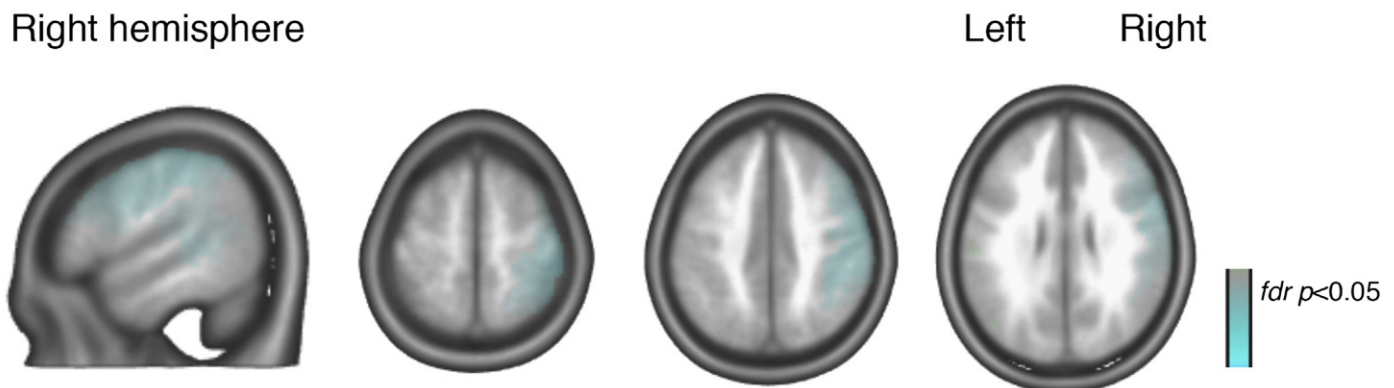
Cerebral source localization of the generators of modulated topographies during the 496–542 ms time window was of special interest, given the fMRI data on the same paradigm. Whole-brain, between-

## Whole-brain statistical contrasts

### rhyme > tone task (496–542 ms)



### tone > rhyme task (804–836 ms)



**Fig. 4.** Cerebral substrates of the effect of selective attention to phonology. Top panel: voxels showing significantly greater current source density ( $fdr p < 0.05$ ) during rhyming versus tone judgments as revealed by whole-brain contrast 496–542 ms following first auditory stimulus. Bottom panel: voxels of the opposite contrast (tone > rhyme judgments) 804–836 ms.

task contrasts revealed that a left-lateralized network of extensive temporal and more posterior cerebral sources was significantly more active while attending to rhymes than to melodies. The current EEG results help identify the time period of the most robust recruitment of the left-lateralized cortical network subserving selective attention to phonology (Yoncheva et al., 2010), as the encoding period for the first stimulus of each trial. This finding reinforces the notion that the observed attentional effects represent modulation of encoding, rather than decision and response properties uniquely associated with the second stimulus. Furthermore, the engagement of the left-lateralized sources coincided with the delivery of the later, rhyme-relevant stimulus portion. This recruitment pattern suggests that the left-lateralized regions deemed involved in phonological processing may be distinct from those involved in attention processes acting on the early P2/N2 components, since these phonological processing regions are active late during stimulus encoding.

The second interval where significant differences in source activation were found was the last distinct segment (804–836 ms) in the inter-stimulus interval. Unlike the pattern of cerebral source involvement in the 496–542 ms window, here only nodes where tone task activations were greater than rhyme task survived the  $fdr p < 0.05$  threshold. Stronger responses for selective attention to melody were observed in extensive bilateral frontal cortex, including the left inferior frontal gyrus and pre-motor areas, the anterior cingulate gyrus, as well as a more focal

right-lateralized cluster extending superior and middle temporal gyri. These cortical regions largely overlap with the network proposed to be involved in the retrieval and anticipation of sound sequences (Leaver et al., 2009; Rauschecker, 2005). Such right-lateralized temporal selective attention effect is also in agreement with the sensitivity of temporal regions to speech versus melodic perception (reviewed in Zatorre, 2003).

Lastly, the rigorous statistical approach applied here is worth reiterating. The converging findings from ERP source localization – which is typically employed in a qualitative manner for illustration of putative cerebral generators – resulted from a data-driven, whole-brain interrogation of each time-point in the trial to determine the time-window of interest, followed by  $t$ -tests at each source node within the brain, and were corrected for multiple comparisons, both temporally and spatially.

The present findings complement an increasingly richer landscape of fMRI investigations examining the relations between auditory and visual language processes. Bitan, Booth and colleagues demonstrated that attention to one type of linguistic information, e.g., phonology, as opposed to another type, e.g., orthography, modulates cortical responses within the language network (Bitan et al., 2005, 2007). Crucially, recent developmental studies have shown that tasks that require active phonological analysis of auditory words activate the ventral visual regions that support reading expertise, and such effects grow stronger as literacy skills develop over childhood and adolescence.



(Booth et al., 2007; Cone et al., 2008). Such findings support the notion that reading acquisition transforms the functional networks that support both written and spoken word processing (Dehaene et al., 2010). Our current work highlights the role of selective attention as a potential driving force for such interactions and temporal dynamics that are both highly dynamic and concurrent with initial encoding of auditory stimuli.

Finally, consideration of potential response interference across tasks provides important insights into the nature of these selective attention effects. Rhyme versus non-rhyme trials had no significant influence on tone judgments; just as, matching versus mismatching tones had no significant influence on rhyme judgments. These reaction time and accuracy results indicate that participants were successful in focusing entirely on the instructed task, and engaging phonological processes exclusively during the rhyme task. Collectively, these behavioral results, the recruitment of the left-lateralized language network specifically during the most relevant stimulus period for phonological analysis, and the underlying “perception for action” CNV component converge to support the conclusion that the current findings reflect selective attention to phonology.

## Conclusions

Here we have demonstrated that selective attention to phonology engaged a left-lateralized network of posterior and extensive temporal regions while the critical rhyme information was encoded within the first word. This goal-directed recruitment of left-lateralized posterior regions included the ventral visual stream, which harbors the visual word form area, whose sensitivity to orthographic stimulus properties and functional contribution to fluent reading skill have been well-established (Cohen et al., 2002; McCandliss et al., 2003) and associated with left-lateralized ERP responses (Maurer et al., 2008). Combining these data-driven ERP findings with the fMRI results on the same paradigm (Yoncheva et al., 2010) suggests the selective involvement of the left mid-fusiform gyrus, which sub-serves literacy, specifically when attention to linguistic information is required in a dynamic fashion. These findings integrate well within a broader theoretical framework that proposes the pivotal role of selective attention in literacy acquisition as a mechanism for integrating emergent phonological skills and reading expertise in left-lateralized language circuits (McCandliss and Yoncheva, 2011). Individual differences in the propensity or ability to focus on phonological information associated with spoken words likely contribute to developmental reading disability (Schlaggar and McCandliss, 2007). Moreover, findings of general selective auditory attention influences on early auditory ERPs across typical literacy development (Coch et al., 2005) and selective language impairment (Stevens et al., 2006), together with recent demonstrations that selective auditory attention can be enhanced through training (Stevens et al., 2008), raise questions regarding the specific nature of selective attention being accessed by these paradigms, modeled on classic dichotic listening. The current paradigm may prove useful in differentiating selective attention based on spatially segregated channels from selective attention to phonological processes per se, thereby providing a closer tie to research on the role of meta-linguistic phonological awareness skills in early literacy acquisition. Illuminating the neural mechanisms that mediate selective attention to phonology will be vital to a better understanding of the interaction between literacy and phonological awareness, a process central to both typical and atypical reading development.

## Acknowledgments

This research was supported by NIMH grants R01DC007694 and 5T32MH067763.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.neuroimage.2014.04.006>.

## References

- Alho, K., Rinne, T., Herron, T.J., Woods, D.L., 2014. Stimulus-dependent activations and attention-related modulations in the auditory cortex: a meta-analysis of fMRI studies. *Hear. Res.* 307, 29–41.
- Astheimer, L.B., Sanders, L.D., 2009. Listeners modulate temporally selective attention during natural speech processing. *Biol. Psychol.* 80, 23–34.
- Berg, P., Scherg, M., 1994. A multiple source approach to the correction of eye artifacts. *Electroencephalogr. Clin. Neurophysiol.* 90, 229–241.
- Bitan, T., Booth, J.R., Choy, J., Burman, D.D., Gitelman, D.R., Mesulam, M.M., 2005. Shifts of effective connectivity within a language network during rhyming and spelling. *J. Neurosci.* 25, 5397–5403.
- Bitan, T., Cheon, J., Lu, D., Burman, D.D., Gitelman, D.R., Mesulam, M.M., Booth, J.R., 2007. Developmental changes in activation and effective connectivity in phonological processing. *Neuroimage* 38, 564–575.
- Bonte, M., Blomert, L., 2004. Developmental changes in ERP correlates of spoken word recognition during early school years: a phonological priming study. *Clin. Neurophysiol.* 115, 409–423.
- Booth, J.R., Cho, S., Burman, D.D., Bitan, T., 2007. Neural correlates of mapping from phonology to orthography in children performing an auditory spelling task. *Dev. Sci.* 10, 441–451.
- Bregman, A.S., 1990. *Auditory Scene Analysis: The Perceptual Organization of Sound*. MIT Press, Cambridge, MA, USA.
- Brem, S., Bach, S., Kucian, K., Guttorm, T.K., Martin, E., Lyytinen, H., Brandeis, D., Richardson, U., 2010. Brain sensitivity to print emerges when children learn letter-speech sound correspondences. *Proc. Natl. Acad. Sci. U. S. A.* 107, 7939–7944.
- Coch, D., Sanders, L.D., Neville, H.J., 2005. An event-related potential study of selective auditory attention in children and adults. *J. Cogn. Neurosci.* 17, 605–622.
- Cohen, L., Lehericy, S., Chochon, F., Lemer, C., Rivaud, S., Dehaene, S., 2002. Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain* 125, 1054–1069.
- Cone, N.E., Burman, D.D., Bitan, T., Bolger, D.J., Booth, J.R., 2008. Developmental changes in brain regions involved in phonological and orthographic processing during spoken language processing. *Neuroimage* 41, 623–635.
- Dehaene, S., Pegado, F., Braga, L.W., Ventura, P., Nunes Filho, G., Jobert, A., Dehaene-Lambertz, G., Kolinsky, R., Morais, J., Cohen, L., 2010. How learning to read changes the cortical networks for vision and language. *Science* 330, 1359–1364.
- Desroches, A.S., Newman, R.L., Joanisse, M.F., 2009. Investigating the time course of spoken word recognition: electrophysiological evidence for the influences of phonological similarity. *J. Cogn. Neurosci.* 21, 1893–1906.
- Ding, N., Simon, J.Z., 2012. Emergence of neural encoding of auditory objects while listening to competing speakers. *Proc. Natl. Acad. Sci. U. S. A.* 109, 11854–11859.
- Dosenbach, N.U., Visscher, K.M., Palmer, E.D., Miezin, F.M., Wenger, K.K., Kang, H.C., Burgund, E.D., Grimes, A.L., Schlaggar, B.L., Petersen, S.E., 2006. A core system for the implementation of task sets. *Neuron* 50, 799–812.
- Dosenbach, N.U., Fair, D.A., Miezin, F.M., Cohen, A.L., Wenger, K.K., Dosenbach, R.A.T., Fox, M.D., Snyder, A.Z., Vincent, J.L., Raichle, M.E., Schlaggar, B.L., Petersen, S.E., 2007. Distinct brain networks for adaptive and stable task control in humans. *Proc. Natl. Acad. Sci. U. S. A.* 104, 11073–11078.
- Dosenbach, N.U., Fair, D.A., Cohen, A.L., Schlaggar, B.L., Petersen, S.E., 2008. A dual-networks architecture of top-down control. *Trends Cogn. Sci.* 12, 99–105.
- Dumay, N., Benraiss, A., Barriol, B., Colin, C., Radeau, M., Besson, M., 2001. Behavioral and electrophysiological study of phonological priming between bisyllabic spoken words. *J. Cogn. Neurosci.* 13, 121–143.
- Efron, B., 2004. Large-scale simultaneous hypothesis testing: the choice of a null hypothesis. *J. Am. Stat. Assoc.* 99, 96–104.
- Efron, B., 2007. Correlation and large-scale simultaneous significance testing. *J. Am. Stat. Assoc.* 102, 93–103.
- Fan, J., Kolster, R., Hajar, J., Suh, M., Knight, R.T., Sarkar, R., McCandliss, B.D., 2007. Response anticipation and response conflict: an event-related potential and functional magnetic resonance imaging study. *J. Neurosci.* 27, 2272–2282.
- Fritz, J.B., Elhilali, M., David, S.V., Shamma, S.A., 2007. Auditory attention—focusing the searchlight on sound. *Curr. Opin. Neurobiol.* 17, 437–455.
- Grave de Peralta Menendez, R., Gonzalez Andino, S., 2002. Comparison of algorithms for the localization of focal sources: evaluation with simulated data and analysis of experimental data. *Int. J. Bioelectromagn.* 4, 1–21.
- Grave de Peralta Menendez, R., Gonzalez Andino, S., Lantz, G., Michel, C.M., Landis, T., 2001. Noninvasive localization of electromagnetic epileptic activity. I. Method descriptions and simulations. *Brain Topogr.* 14, 131–137.
- Grave de Peralta Menendez, R., Murray, M.M., Michel, C.M., Martuzzi, R., Gonzalez Andino, S., 2004. Electrical neuroimaging based on biophysical constraints. *Neuroimage* 21, 527–539.
- Hamano, T., Luders, H.O., Ikeda, A., Collura, T.F., Comair, Y.G., Shibasaki, H., 1997. The cortical generators of the contingent negative variation in humans: a study with subdural electrodes. *Electroencephalogr. Clin. Neurophysiol.* 104, 257–268.
- Hillyard, S.A., Hink, R.F., Schwent, V.L., Picton, T.W., 1973. Electrical signs of selective attention in the human brain. *Science* 182, 177–180.

- Hillyard, S.A., Vogel, E.K., Luck, S.J., 1998. Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 353, 1257–1270.
- Hink, R.F., Hillyard, S.A., 1976. Auditory evoked potentials during selective listening to dichotic speech messages. *Percept. Psychophys.* 20, 236–242.
- Hink, R.F., Hillyard, S.A., Benson, P.J., 1978. Event-related brain potentials and selective auditory attention to acoustic and phonetic cues. *Biol. Psychol.* 6, 1–16.
- Joanisse, M.F., Zevin, J.D., McCandliss, B.D., 2007. Brain mechanisms implicated in the preattentive categorization of speech sounds revealed using fMRI and a short-interval habituation trial paradigm. *Cerebral Cortex* 17 (9), 2084–2093.
- Leaver, A.M., Van Lare, J., Zielinski, B., Halpern, A.R., Rauschecker, J.P., 2009. Brain activation during anticipation of sound sequences. *J. Neurosci.* 29, 2477–2485.
- Lehmann, D., Skrandies, W., 1980. Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalogr. Clin. Neurophysiol.* 48, 609–621.
- Luu, P., Ferree, T., 2000. Determination of the Geodesic Sensor Nets' average electrode positions and their 10–10 international equivalents.
- Maurer, U., Rossion, B., McCandliss, B.D., 2008. Category specificity in early perception: face and word n170 responses differ in both lateralization and habituation properties. *Front. Hum. Neurosci.* 2, 18.
- McCandliss, B.D., Yoncheva, Y., 2011. Integration of left-lateralized neural systems supporting skilled reading. In: Benasich, A.A., Fitch, R.A. (Eds.), *Developmental Dyslexia: Early Precursors, Neurobehavioral Markers and Biological Substrates* (The Extraordinary Brain Series). Brookes Publishing Co., Baltimore, MD, pp. 325–339.
- McCandliss, B.D., Cohen, L., Dehaene, S., 2003. The visual word form area: expertise for reading in the fusiform gyrus. *Trends Cogn. Sci.* 7, 293–299.
- Mesgarani, N., Chang, E.F., 2012. Selective cortical representation of attended speaker in multi-talker speech perception. *Nature* 485, 233–236.
- Michel, C.M., 2009. *Electrical Neuroimaging*. Cambridge University Press, Cambridge.
- Michel, C.M., Murray, M.M., Lantz, G., Gonzalez, S., Spinelli, L., Grave de Peralta, R., 2004. EEG source imaging. *Clin. Neurophysiol.* 115, 2195–2222.
- Nagai, Y., Critchley, H.D., Featherstone, E., Fenwick, P.B., Trimble, M.R., Dolan, R.J., 2004. Brain activity relating to the contingent negative variation: an fMRI investigation. *Neuroimage* 21, 1232–1241.
- Newman, R.L., Connolly, J.F., 2009. Electrophysiological markers of pre-lexical speech processing: evidence for bottom-up and top-down effects on spoken word processing. *Biol. Psychol.* 80, 114–121.
- Pascual-Marqui, R.D., Michel, C.M., Lehmann, D., 1995. Segmentation of brain electrical activity into microstates: model estimation and validation. *IEEE Trans. Biomed. Eng.* 42, 658–665.
- Picton, T.W., Hillyard, S.A., 1974. Human auditory evoked potentials. II. Effects of attention. *Electroencephalogr. Clin. Neurophysiol.* 36, 191–199.
- Power, A.J., Foxe, J.J., Forde, E.-J., Reilly, R.B., Lalor, E.C., 2012. At what time is the cocktail party? A late locus of selective attention to natural speech. *Eur. J. Neurosci.* 35, 1497–1503.
- Praamstra, P., Stegeman, D.F., 1993. Phonological effects on the auditory N400 event-related brain potential. *Cogn. Brain Res.* 1, 73–86.
- Praamstra, P., Meyer, A., Levett, W.J.M., 1994. Neurophysiological manifestations of phonological processing: latency variation of negative ERP component time-locked to phonological mismatch. *J. Cogn. Neurosci.* 6, 204–219.
- R Development Core Team, 2007. *R: A language and environment for statistical computing*. Rauschecker, J.P., 2005. Neural encoding and retrieval of sound sequences. *Ann. N. Y. Acad. Sci.* 1060, 125–135.
- Reynolds, J.R., West, R., Braver, T., 2009. Distinct neural circuits support transient and sustained processes in prospective memory and working memory. *Cereb. Cortex* 19, 1208–1221.
- Sanders, L.D., Newport, E.L., Neville, H.J., 2002. Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. *Nat. Neurosci.* 5, 700–703.
- Schlaggar, B.L., McCandliss, B.D., 2007. Development of neural systems for reading. *Annu. Rev. Neurosci.* 30, 475–503.
- Schroeder, C.E., Lakatos, P., Kajikawa, Y., Partan, S., Puce, A., 2008. Neuronal oscillations and visual amplification of speech. *Trends Cogn. Sci.* 12, 106–113.
- Schroeder, C.E., Wilson, D.A., Radman, T., Scharfman, H., Lakatos, P., 2010. Dynamics of Active Sensing and perceptual selection. *Curr. Opin. Neurobiol.* 20, 172–176.
- Spinelli, L., Andino, S.G., Lantz, G., Seeck, M., Michel, C.M., 2000. Electromagnetic inverse solutions in anatomically constrained spherical head models. *Brain Topogr.* 13, 115–125.
- Sridharan, D., Levitin, D.J., Chafe, C.H., Berger, J., Menon, V., 2007. Neural dynamics of event segmentation in music: converging evidence for dissociable ventral and dorsal networks. *Neuron* 55, 521–532.
- Stevens, C., Sanders, L.D., Neville, H.J., 2006. Neurophysiological evidence for selective auditory attention deficits in children with specific language impairment. *Brain Res.* 1111, 143–152.
- Stevens, C., Fanning, J., Coch, D., Sanders, L.D., Neville, H.J., 2008. Neural mechanisms of selective auditory attention are enhanced by computerized training: electrophysiological evidence from language-impaired and typically developing children. *Brain Res.* 1205, 55–69.
- Strik, W.K., Fallgatter, A.J., Brandeis, D., Pascual-Marqui, R.D., 1998. Three-dimensional tomography of event-related potentials during response inhibition: evidence for phasic frontal lobe activation. *Electroencephalogr. Clin. Neurophysiol.* 108, 406–413.
- Strimmer, K., 2008a. A unified approach to false discovery rate estimation. *BMC Bioinforma.* 9, 303.
- Strimmer, K., 2008b. *fdrttool: a versatile R package for estimating local and tail area-based false discovery rates*. *Bioinformatics* 24, 1461–1462.
- Tucker, D.M., 1993. Spatial sampling of head electrical fields: the geodesic sensor net. *Electroencephalogr. Clin. Neurophysiol.* 87, 154–163.
- Woodcock, R.W., McGrew, K., Mather, N., 2001. *Woodcock-Johnson III*. Riverside Publishing, Itasca, IL, USA.
- Woods, D.L., Hillyard, S.A., Hansen, J.C., 1984. Event-related brain potentials reveal similar attentional mechanisms during selective listening and shadowing. *J. Exp. Psychol. Hum. Percept. Perform.* 10, 761–777.
- Yoncheva, Y., Zevin, J., Maurer, U., McCandliss, B.D., 2010. Auditory selective attention to speech modulates activity in the visual word form area. *Cereb. Cortex* 20, 622–632.
- Yoncheva, Y., Maurer, U., Zevin, J.D., McCandliss, B.D., 2013. Effects of rhyme and spelling patterns on auditory word ERPs depend on selective attention to phonology. *Brain Lang.* 124, 238–243.
- Zatorre, R.J., 2003. Sound analysis in auditory cortex. *Trends Neurosci.* 26, 229–230.
- Ziegler, J.C., Goswami, U., 2005. Reading acquisition, developmental dyslexia and skilled reading across languages: a psycholinguistic grain size theory. *Psychol. Bull.* 131, 3–29.